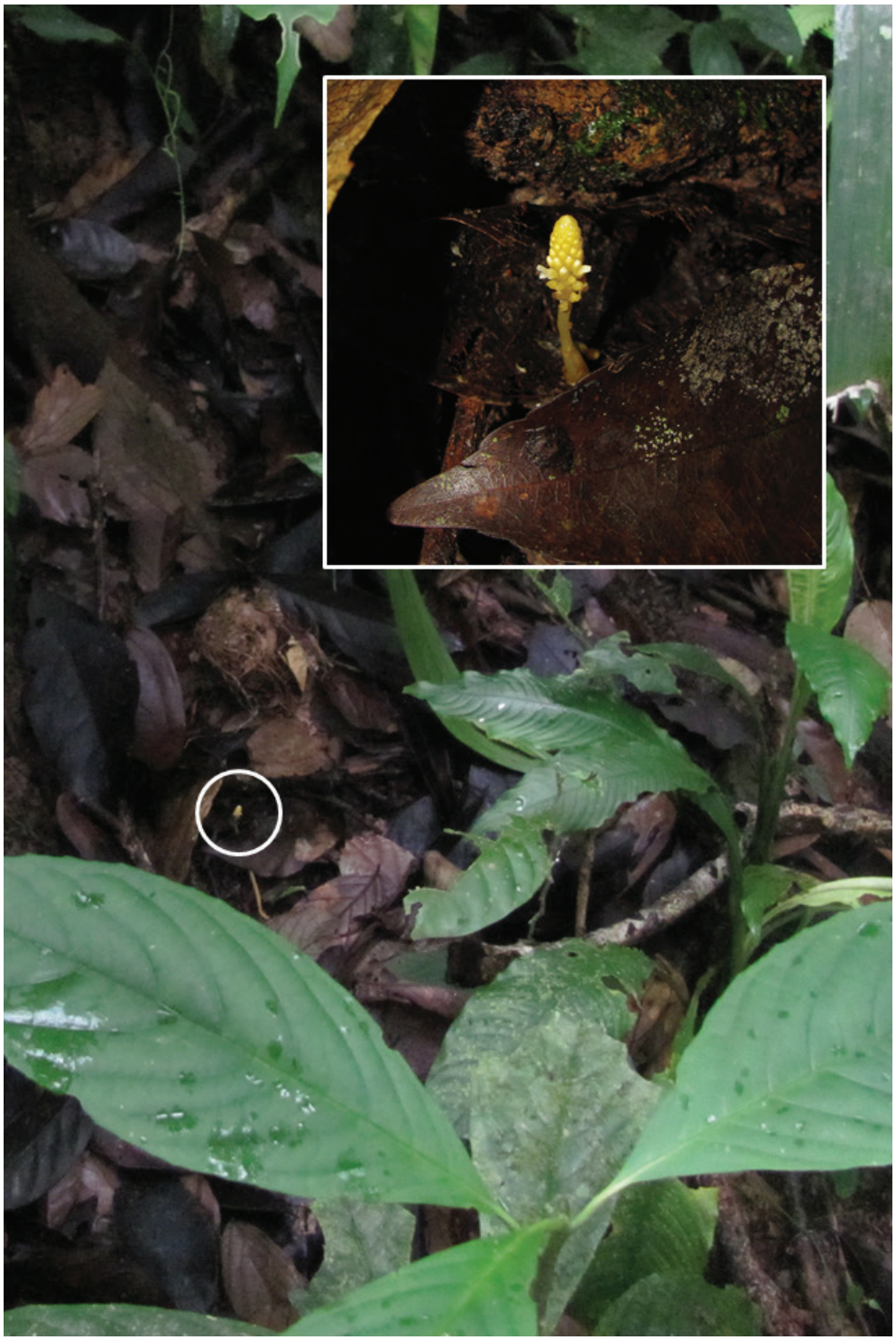
Antiquities of the rainforest: evolution of mycoheterotrophic angiosperms growing on Glomeromycota

Mennes, C.B.

Chapter 4

Evolution of mycoheterotrophy in Polygalaceae: The case of *Epirixanthes*

Next page: *Epirixanthes pallida* growing in Kinabalu National Park, Sabah, Malaysian Borneo



Evolution of mycoheterotrophy in Polygalaceae: The case of *Epirixanthes*

Constantijn B. Mennes[1], Michelangelo S. Moerland[1], Magnus Rath[3], Erik F. Smets[1,2], Vincent S. F. T. Merckx[1]

[1]Naturalis Biodiversity Center, Leiden University, P.O. Box 9517, Leiden, the Netherlands

[2]Section Ecology, Evolution and Biodiversity Conservation, KU Leuven, Kasteelpark Arenberg 31, P.O. Box 2437, BE-3001 Leuven, Belgium

[3]Spezielle Botanik und Mykologie, Fachbereich Biologie, Philipps-Universität Marburg 35032 Marburg, Germany

Published in *American Journal of Botany* 102: 1-11 (2015).

4.1. Abstract

Premise of the study: The mycoheterotrophic lifestyle has enabled some plant lineages to obtain carbon from their mycorrhizal symbionts. The mycoheterotrophic genus *Epirixanthes* (Polygalaceae) consists of six species from tropical Asia. Although it is probably closely related to the chlorophyllous genus *Salomonia* and linked to arbuscular mycorrhizal fungi, lack of DNA sequence data has thus far prevented these hypotheses from being tested. Therefore, the evolutionary history of *Epirixanthes* remains largely unknown.

Methods: We reconstructed the phylogenetic relationships of *Epirixanthes* based on nuclear ITS and plastid *matK* data. Divergence times were inferred using a Bayesian relaxed clock approach, and we phylogenetically analyzed its mycorrhizal symbionts. We furthermore assigned these symbionts to operational taxonomic units, compared them with symbionts of other Polygalaceae, and measured their phylogenetic diversity.

Key results: We found that *Epirixanthes* is placed in tribe Polygaleae as sister to *Salomonia*. *Epirixanthes* has a Miocene-Oligocene stem age and grows exclusively in symbiosis with fungi of Glomeraceae. *Salomonia* and some *Polygala* species are linked to both Glomeraceae and Acaulosporaceae, resulting in higher phylogenetic diversity values. The majority of the symbionts of *Epirixanthes* are not found in *Salomonia* or *Polygala*, although a few shared fungal taxa are found.

Conclusions: Epirixanthes forms a relatively young mycoheterotrophic lineage. The Oligocene-Miocene origin suggests its evolution was influenced by the environmental dynamics in Southeast Asia during this time. Although comparison of fungi from *Epirixanthes* with those from *Salomonia* and *Polygala* suggests some specialization, many other mycoheterotrophic plants are linked to a more narrow set of Glomeraceae.

Keywords: arbuscular mycorrhizal fungi; *Epirixanthes*; Glomeraceae; mycoheterotrophy; Polygalaceae.

4.2. Introduction

Several plant lineages have developed ways to obtain carbon from their mycorrhizal symbionts, a mode of life known as mycoheterotrophy (Leake, 1994). This symbiotic interaction evolved from the common mutualistic interactions between plants and mycorrhizal fungi. While most autotrophic plants are mycorrhizal generalists and interact with multiple fungi at the same time, many mycoheterotrophs are specialists and associate with a narrow range of fungi only (Bidartondo, 2005). The majority of mycoheterotrophic lineages are monocots, both in species diversity and number of independent origins (Merckx et al., 2013a). Within the eudicots, the mycoheterotrophic lifestyle only occurs in three families: Gentianaceae, Ericaceae, and Polygalaceae (Merckx and Freudenstein, 2010). Whereas the phylogenetic position of the mycoheterotrophic Gentianaceae, and Ericaceae, and the identity of their fungal symbionts were studied recently (Bidartondo and Bruns, 2002; Merckx et al., 2012, 2013c), such data are lacking for the mycoheterotrophic representatives of Polygalaceae (Fabales). Furthermore, Polygalaceae are the last angiosperm family for which no sequence data are available for its mycoheterotrophic representatives.

The genus *Epirixanthes* Blume is the only fully mycoheterotrophic lineage in Polygalaceae and contains the only mycoheterotrophic representatives of the rosids (Stevens, 2001 onward; Merckx et al., 2013a). *Epirixanthes* consists of six species (*Epirixanthes compressa* Pendry, *E. cylindrica* Blume, *E. elongata* Blume, *E. kinabaluensis* Wendt, *E. pallida* Wendt, and *E. papuana* J.J.Smith), all from tropical Asia (Van der Meijden, 1988; Pendry, 2010). The species' distributions overlap and range from eastern India and China throughout Malesia to the Solomon Islands: *Epirixanthes compressa* was recently described from Thailand (Pendry, 2010); *Epirixanthes elongata* has a wide distribution and is found in eastern India, southern Myanmar, northern Vietnam, southern China, Peninsular Malaysia, Sumatra, western Java, Borneo, and Moluccas (Van der Meijden, 1988); *Epirixanthes papuana* grows in northern Sumatra, western Java, Borneo, Philippines, Moluccas, New Guinea, and the Solomon Islands (Van der Meijden, 1988); *Epirixanthes cylindrica* occurs in Myanmar, Sumatra, western Java, Borneo, and New Guinea (Van der Meijden, 1988); *Epirixanthes kinabaluensis* is only found in Borneo and Sumatra (Van der Meijden, 1988); and *Epirixanthes pallida* has a relatively narrow distribution restricted to Borneo and Sulawesi (Van der Meijden, 1988). Borneo is the center of diversity of *Epirixanthes*, with five of the six species present. Some authors (e.g., Chodat, 1896; Hutchinson, 1967) suggested *Epirixanthes* to be congeneric with *Salomonia* Lour., a genus of two (Van der Meijden, 1988) to five (Pendry, 2001) species of autotrophic herbs. Van der Meijden (1988) separated these two genera on the basis of differences in several morphological characters (*Epirixanthes* vs. *Salomonia*: mycoheterotrophic vs. autotrophic; indehiscent vs. dehiscent fruits; disk present vs. absent; straight or short style with anthers not enclosing the stigma vs. strongly curved style with anthers tightly enclosing the stigma). The classification in

two separate genera is used by many subsequent authors (Eriksen, 1993a, b; Forest et al., 2007; Imhof, 2007; Bello et al., 2009, 2010, 2012; Pendry, 2010), though it is not used in the review on mycoheterotrophy by Leake (1994). By identifying a few morphological synapomorphies for the tribe Polygaleae (arrested lateral petals) and the genera *Epirixanthes* and *Salomonia* (three antesepalous stamen primordia), the results of Bello et al. (2012) indicated a close relationship between the genera, as well as placement in the tribe Polygaleae. *Salomonia* has been consistently placed in Polygaleae in molecular phylogenetic reconstructions of Polygalaceae (Persson, 2001; Forest et al., 2007; Bello et al., 2009, 2012). The lack of sequence data of *Epirixanthes,* however, has prevented testing its proposed close relationship with *Salomonia.*

Fungal symbionts of *Epirixanthes* have been identified as arbuscular mycorrhizal fungi (AMF) of the *Paris* type (Imhof, 2007), but no family-level identification based on DNA sequence data has been carried out thus far. Imhof (2007) suggested a degree of adaptation to these fungi and reported similarities in AMF features with unrelated mycoheterotrophic plants and with published drawings of mycorrhizal symbionts of *Polygala amara* (Marcuse, 1902). Within Polygalaceae, the chlorophyllous genus *Polygala* was found to grow in symbiosis with several clades of AMF of the family Glomeraceae, based on DNA sequence data (Rath et al., 2013). Moreover, a single fungal sequence from Acaulosporaceae was extracted from *P. myrtifolia.* This confirms the expected low specificity of the arbuscular mycorrhizal symbiosis in autrotrophic species. For mycoheterotrophic groups, however, specialization on narrow lineages of AMF is often observed and thus expected (Merckx et al., 2009b, 2010b). *Petrosavia sakuraii*, for example, is found to grow in symbiosis with a small subset of the mycorrhizal symbionts found in the sister group of *Petrosavia*, *Japonolirion osense* (Yamato et al., 2014). However, some groups (i.e., *Sciaphila* [Triuridaceae] and *Campylosiphon* [Burmanniaceae]) appear to be less specialized as they were found to grow in symbiosis with a relatively wide range of AMF (Merckx et al., 2012). These different patterns of specialization raise the question of whether *Epirixanthes* is specialized on narrow lineages of AMF as compared with its autotrophic relatives and other mycoheterotrophic lineages.

The major goal of this study is to evaluate the phylogenetic position of *Epirixanthes* in Polygalaceae and to infer the divergence age of the group. Subsequently, fungal symbionts are identified and analyzed in a phylogenetic framework of AMF. We use this framework to compare the fungal specificity between *Epirixanthes* and its close relatives in *Salomonia* and *Polygala* by measuring the phylogenetic diversity (PD; Faith, 1992) of the associated AMF. The results were compared with published insights on AMF of Polygalaceae (Imhof, 2007; Rath et al., 2013) as well as to recent studies of mycoheterotrophic specificity (Hynson and Bruns, 2009; Merckx et al., 2009b, 2010b, 2012; Roy et al., 2009).

4.3. Materials and Methods

4.3.1. Taxon information

Plant material and root tips containing endophytic fungal material of Epirixanthes were collected at three localities in Kinabalu National Park and Crocker Range National Park in Sabah, Malaysian Borneo. Collections were made during the collaborative 2012 Kinabalu-Crocker Range scientific expedition organized by Sabah Parks (Malaysia) and Naturalis Biodiversity Center (the Netherlands). Plant identifications followed the key of Van der Meijden (1988). Flowering specimens of the species *Epirixanthes pallida, E. papuana, E. kinabaluensis,* and *E. elongata* were collected. Additional Polygalaceae sequence data were obtained from the collections of Naturalis (L), Royal Botanic Gardens, Kew (K), and GenBank. Appendix A.1 includes a full list of included specimens. Two species of *Salomonia* were collected at two localities in Thailand, identified using the key by Pendry (2001), cleaned with water and dried with silica gel for subsequent molecular analyses of the fungal symbionts. Two individuals of *Salomonia longiciliata* (i1 and i2) were collected in 2011 in a dipterocarp forest in Saloung subdistrict, Mae Rim district in Chiang Mai province. Two individuals of *Salomonia cantoniensis* (i1 and i2) were collected in a dipterocarp forest near Ban Moung Soi village, Pai district in Mae Hong Son province. All plants of *Salomonia* were flowering and additionally bore mature fruits. Additional fungal sequence data represented virtual taxa (i.e., fungal phylogroups based on 97% sequence similarity and bootstrap support) from the Maarj*AM* database (Öpik et al., 2010), as well as mycorrhizal symbionts of *Polygala* from Rath et al. (2013).

4.3.2. Amplification of plant loci

DNA was extracted using the Qiagen DNEasy Plant Mini Kit (Qiagen, Venlo, Netherlands), according to the manufacturer's protocol. The nuclear internal transcribed spacer (ITS) region was amplified, as well as the plastid maturase K gene (*matK*). For ITS, primers ITS1 (White et al., 1990) and ITS4 (Wojciechowski et al., 1993) were used. The reaction mixture contained 1 µL of undiluted DNA, 1 µL of every primer (0.2 µM), 2.5 µL dNTP, 2.5 µl 10× NH4 reaction buffer (Bioline, Luckenwalde, Germany), 1 µL $MgCl_2$ (2 µM) and 0.2 µL BioTaq DNA Polymerase (Qiagen) and 15.8 µL milliQ water, to a total of 25 µL. The PCR procedure consisted of 4–5 min of initial denaturation at 95°C followed by 30 cycles of 30 s at 95°C for denaturation, 30 s at 44°C for annealing and 30 s at 72°C for primer extension, followed by a final extension of 7 min at 72°C. The PCR for samples from the Kew DNA Bank (see Appendix A.1) consisted of 4 min of initial denaturation at 94°C; followed by 40 cycles of 45 s at 94°C for denaturation, 60 s at 48°C for annealing, and 90 s at 72°C for primer extension; and a final extension of 7 min at 72°C. For *matK*, primers *trnK*685F (Hu et al., 2000) and *matK*1777R (Hu et al., 2000) were

used at a concentration of 1 µM. The remaining reaction mixture was identical to the one used for ITS. The PCR procedure was based on Hu et al. (2000) and Sirichamorn et al. (2012) and consisted of 4 min of initial denaturation at 94°C; followed by 40 cycles of 45 s at 94°C for denaturation, 90 s at 48°C for annealing, and 90 s at 72°C for primer extension; and a final extension of 7 min at 72°C.

4.3.3. Amplification and cloning of AMF loci

Fungal DNA was extracted from *Epirixanthes* root tips stored in 2% cetyltrimethyl ammonium bromide (CTAB) buffer. Extraction followed the DNA Plant Mini Kit (Qiagen) according to the manufacturer's protocol. A 500-bp fragment of the nuclear 18S rDNA region was amplified using the primer combination NS31 (Simon, 1996) and AM1 (Helgason et al., 1998). The reaction mixture was identical to the one used for PCR of plant DNA as described above, using primer concentrations of 5 µM. The PCR procedure consisted of 5 min of initial denaturation at 94°C; 30 cycles of 30 s at 94°C for denaturation, 30 s at 57°C for annealing, and 1 min at 72°C for primer extension; and a final extension of 7 min at 72°C.

Eight PCR sequencing products of poor quality (<50% high-quality bases) were selected for cloning. Four of those were amplified using primer combination NS31 and AM1, and four were amplified using primer combination NS1 (White et al., 1990) and EF3 (Smit et al., 1999). The first four were amplified as described above; the last four were amplified using internal sequencing primers NS2, NS3, NS4, NS5 (White et al., 1990). The PCR procedure of this primer combination consisted of 2 min of initial denaturation at 94°C; 35 cycles of 40 s at 94°C for denaturation, 30 s at 54°C for annealing, and 105 s at 72°C for primer extension; and a final extension of 7 min at 72°C. The resulting PCR products were cloned using the TOPO TA Cloning Kit for sequencing, with One Shot Mach 1 T1R Chemically Competent *E. coli*, according to the manufacturer's protocol (Invitrogen, Life Technologies, San Diego, California, USA). Eight white colonies were selected from each plate corresponding to the cloned PCR product. These colonies (~0.3 µL) were used as the DNA template in a PCR reaction mixture, consisting of 0.5 µL of every vector-associated primer (M13F and M13R), 1 µL dNTP (Qiagen, 2.5 mM), 3 µL CoralLoad reaction buffer (Qiagen), 2 µL MgCl2 (25 mM), and 0.2 µL BioTaq DNA polymerase (Qiagen, 5 U/µL), and 23.1 µL milliQ water, to a total of roughly 30 µL. The PCR procedure consisted of 3 min of initial denaturation at 95°C; 35 cycles of 20 s at 95°C for denaturation, 20 s at 50°C for annealing, and 2 min at 72°C for primer extension; and a final extension of 10 min at 72°C.

Fungal DNA from the species of *Salomonia* was extracted and amplified as follows. From each plant of the two species a single root fragment (1–3 cm) was used for molecular identification of the mycorrhizal symbionts. The PCR procedure for the 18S rDNA region of the fungal symbionts basically followed the DNA extraction

protocol described in Rath et al. (2013), although here we used dried root material instead of CTAB fixed material. Amplification of fungal DNA was done using a nested PCR approach. The first primer combination of GeoA1 and ART4 (Schwarzott and Schüßler, 2001) was used for initial amplification of 18S rDNA to obtain an initial PCR product, which increases the quality and quantity of the subsequent PCR reactions. These subsequent reactions were conducted by nested PCR using the AMF-specific primer combination AM1 and NS31 (Simon, 1996; Helgason et al., 1998). The first PCR mixture consisted of 15.5 µL sterile double-distilled water, 2.5 µL 10× PCR buffer, 1 µL of each primer (10 µM), 2.5 µL of dNTP (5 mM), 0.5 µL Pfu polymerase (Promega, Madison, Wisconsin, USA), and 2.0 µL DNA extract as a template. The PCR procedure, adapted from Schwarzott and Schüßler (2001), consisted of 2 min for initial denaturation at 95°C followed by 30 cycles of 30 s at 95°C for denaturation, 30 s at 60°C for annealing and 4 min at 73°C for primer extension, followed by a final extension of 10 min at 73°C. The second PCR mixture consisted of 35.5 µL sterile double-distilled water, 5 µL 10× PCR buffer, 2 µL of each primer (10 µmol/L), 2.5 µL of dNTP (5 mM), 0.5 µL Pfu polymerase, and 2.5 µL of the amplification of the first PCR as a template. The PCR procedure consisted of 2 min for initial denaturation at 95°C; 35 cycles of 30 s at 95°C for denaturation, 30 s at 60°C for annealing, and 1.5 min at 73°C for primer extension; and a final extension of 5 min at 73°C. In case of only weak amplification, up to four additional nested PCR reactions were done using the primer combination AM 1 and NS31, and the resulting products were pooled. Resulting sequences were compared with sequences from GenBank using the BLAST algorithm and selected for direct further analysis if high similarity to AMF could be observed. In only two cases (S. longiciliata i2 and S. cantoniensis i1), initial sequencing of the PCR product with primer AM1 resulted in high sequence quality, which allowed immediate use for the alignment. The remaining PCR products of *S. longiciliata* i1 and i2 and *S. cantoniensis* i2 were used for cloning and subsequent sequencing (see Rath et al. [2013] for detailed information on DNA processing).

4.3.4. Phylogenetic reconstruction of plant data

Sequences were assembled, aligned, and manually edited using the program Geneious Pro version 6.1.7 (Drummond et al., 2013). Sequences were aligned using the program MUSCLE (Edgar, 2004) method as incorporated in the program Geneious Pro (default settings). Appendix S2a (see Supplemental Data with the online version of this article) contains the alignment of the combined plant data set. The combined plant data set was analyzed using a model-based maximum likelihood (ML) as optimality criterion, as well as model-based Bayesian posterior probabilities (PP). Both plant data sets were also analyzed separately using these methods. The fungal data set was analyzed using only the ML approach, supplemented by additional diversity measurements (see further). The best substitution model for each data set under the Akaike information criterion (AIC) was found using the program jModeltest

version 2.1.4 (Guindon and Gascuel, 2003; Darriba et al., 2012). For the ITS data set, the GTR+I+G substitution model was selected. For the *matK* data set, TVM+G was suggested as best fit, but GTR+G was selected as the best model incorporated in the phylogenetic analysis packages. The complete data set was partitioned into two subsets, each corresponding to the two genetic regions and analyzed with the selected substitution model. The ML analyses were conducted using the program Garli version 2.0 (Zwickl, 2006), supplemented by the bootstrap support method (BS; Felsenstein, 1985) to evaluate branch support (500 replications). Majority rule consensus trees were constructed using the Tree Builder plugin in Geneious Pro (Drummond et al., 2013). The Bayesian inference (BI) analyses were carried out using MrBayes version 3.2.2 (Ronquist et al., 2012). The Markov chain Monte Carlo (MCMC; Geyer, 1991) analyses were performed with two runs of four chains and terminated after 4 000 000 generations if the standard deviation of split frequencies was 0.01 or lower. A burn-in of 25% was discarded. Results were evaluated using the program Tracer version 1.6 (Rambaut et al., 2013), by checking the distribution and stability of the likelihood values (normally distributed values were considered sufficient). The effective sample size (ESS) was also evaluated; a minimum value of 200 was considered sufficient. Clades were considered well supported when receiving over 85% BS or 95% PP. The resulting trees from the BI analyses were visualized using FigTree version 1.4 (Rambaut, 2012) and presented with both PP values and BS values resulting from the ML analyses plotted on the nodes. Only support values at the genus level were given, although additional values were given for the species of *Epirixanthes* and for some clades of *Polygala*.

4.3.5. Divergence time estimation

Divergence times of the taxa in Polygalaceae were estimated by using an uncorrelated log normal relaxed clock method using BEAST v2.1.3 (Bouckaert et al., 2014). The ITS and *matK* data sets used in the phylogenetic analyses were used for this purpose. Only a single specimen per species was used, resulting in a data set containing 31 Polygalaceae taxa. *Xanthophyllum* was set as outgroup (Forest et al., 2007). The data set was partitioned in two subsets, corresponding to the two loci. The ITS data set was analyzed using the GTR+I+G substitution model and the *matK* data set was analyzed using the GTR+G substitution model, as selected with jModeltest version 2.1.4 (Guindon and Gascuel, 2003; Darriba et al., 2012) (see section "Phylogenetic reconstruction of plant data"). The Yule birth rate prior (Gernhard, 2008) was set with a flat prior on the birth rate. A single tree model and a single clock model were used. For the clock model, an exponential prior was set for both the mean (ucld.mean; mean = 10.0) and the standard deviation (ucld.stdev; mean= 0.33). The mean substitution rates of both data sets were fixed. Only limited calibration information was available for Polygalaceae. A secondary calibration point prior was set on the root of the tree (i.e., the crown node of Polygalaceae). This prior was modeled as normal distribution with a mean of 60.4 Ma and a standard deviation of 10.0 Ma

and was based on the estimated crown age of Polygalaceae resulting from Forest et al. (2007). All other prior settings were kept to the default settings. Approximation of posterior distributions was done using the Markov chain Monte Carlo (MCMC; Geyer, 1991) procedure and was run with a chain length of 50 000 000 generations. A single run was conducted, and one tree in every 5000 generations was sampled, discarding a burn-in of 10%. The maximum clade credibility tree was constructed using the program TreeAnnotator v1.7.5 (Drummond et al., 2012). Results were visualized using the program Tracer v1.6 (Rambaut et al., 2013) by evaluating ESS values for parameters (see section *Phylogenetic reconstruction of plant data*) as well as the stability and distribution of the obtained likelihood scores (the prior distribution of the scores was considered sufficient). Finally, the resulting tree was visualized using FigTree v1.4 (Rambaut, 2012).

4.3.6. Phylogenetic reconstruction and diversity measurements of AMF

The diversity of arbuscular mycorrhizal symbionts of *Epirixanthes* was assessed by (1) assigning the 500-bp fragment of 18S rDNA of all AMF sequences to operational taxonomic units (OTUs) based on 97% sequence similarity, which is the same percentage as used in defining the virtual taxa in the Maarj*AM* database, although these were also based on bootstrap support values (Öpik et al., 2010), (2) reconstructing the phylogeny of these AMF in a Glomeromycota framework using the Maarj*AM* database and (3) by measuring Faith's phylogenetic diversity (PD) values (Faith, 1992) for the AMF of *Epirixanthes* and comparing them with PD values of AMF of *Salomonia* and *Polygala*. First, an alignment was made using the MUSCLE (Edgar, 2004) method, containing fungal sequences from *Epirixanthes*, *Salomonia*, and *Polygala*, as well as virtual taxa from the Maarj*AM* database (Öpik et al., 2010) representing Glomeraceae, Acaulosporaceae, Gigasporaceae, Diversisporaceae, Claroideoglomerales, and Paraglomeraceae. This alignment was manually trimmed. Appendix S2b (see the online Supplemental Data with the published version of this chapter) contains the alignment of the fungal data set. A subset containing the AMF from *Epirixanthes*, *Salomonia*, and *Polygala* was extracted. The OTUs were assigned using the UPARSE-OTU algorithm (Edgar, 2013) as implemented in Usearch version 7.0 (Edgar, 2010), discarding full-length identical sequences. Chimeric sequences were detected automatically by the UPARSE-OUT algorithm (Edgar, 2013), as sequences that could not be assigned a single OTU. Phylogenetic analysis of the 18S rDNA sequence data from the AMF was similar to the ML analysis of the plant material as described above. The substitution model GTR+I+G was selected as best fit. A highest likelihood ML tree was reconstructed, and 800 bootstrap replicates were performed using the Message Passing Interface (MPI) version of Garli. Support values (60% BS and higher) were given for clades containing either fungal symbionts of *Epirixanthes*, *Salomonia*, or *Polygala*. To calculate PD values of the fungal communities in the roots, the resulting ML tree was combined with a "community data matrix", which linked each amplified fungal sequence to its host

species. The Picante package (Kembel et al., 2010) as incorporated in R (R Core Team, 2014) was used to calculate PD values of each species. The standardized effect sizes (SES) of the PD values were furthermore calculated to correct for the sampling size, by testing the extent to which communities of fungi are more (or less) related to each other than expected based on randomization. This randomization was performed using the "richness" null model as incorporated in the Picante package (Kembel et al., 2010), which randomized abundances of fungi per plant species, while maintaining "species richness" and phylogenetic relations of these fungi.

4.4. Results

4.4.1. Phylogeny of *Epirixanthes*

The phylogeny based on the combined nuclear ITS (805 bp) and chloroplast *matK* (1160 bp) data set indicates *Epirixanthes* is monophyletic (100% BS and 100% PP) and sister to the monophyletic genus *Salomonia* (100% BS and 100% PP) at the

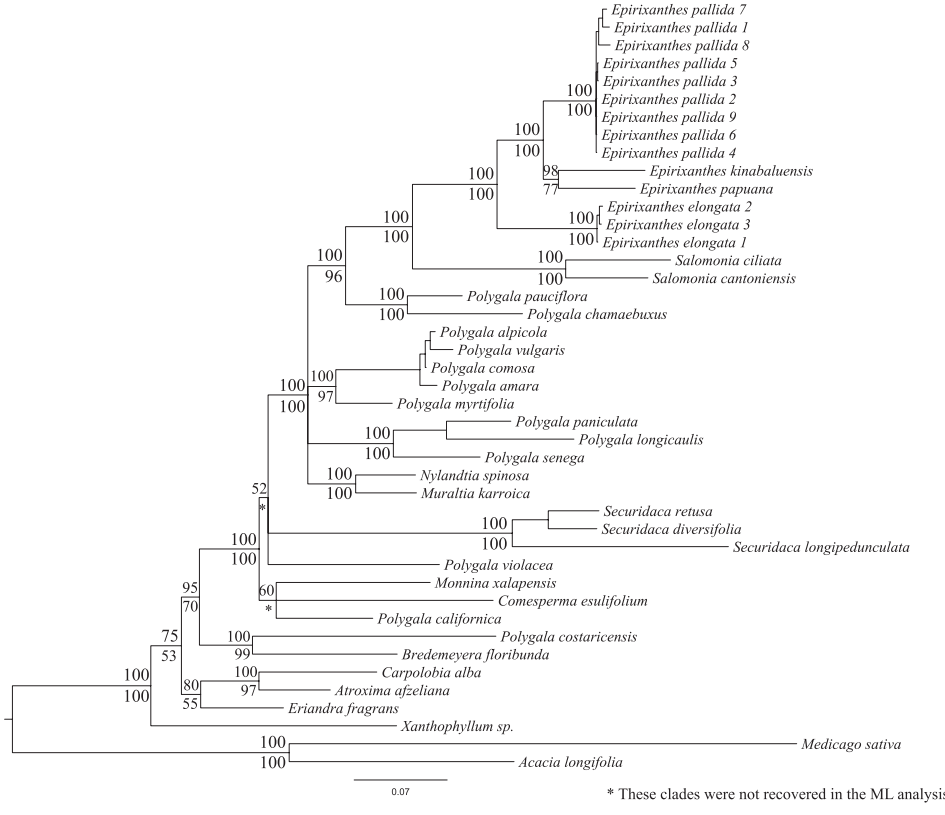


Fig. 13. Phylogeny of Polygalaceae based on a Bayesian Inference (BI) analysis of the combined ITS and *matK* datasets. Values above branches are posterior probabilities resulting from the BI analysis; those below branches are bootstrap percentages resulting from the maximum likelihood (ML) analysis. The scale bar indicates the number of support substitutions per site.

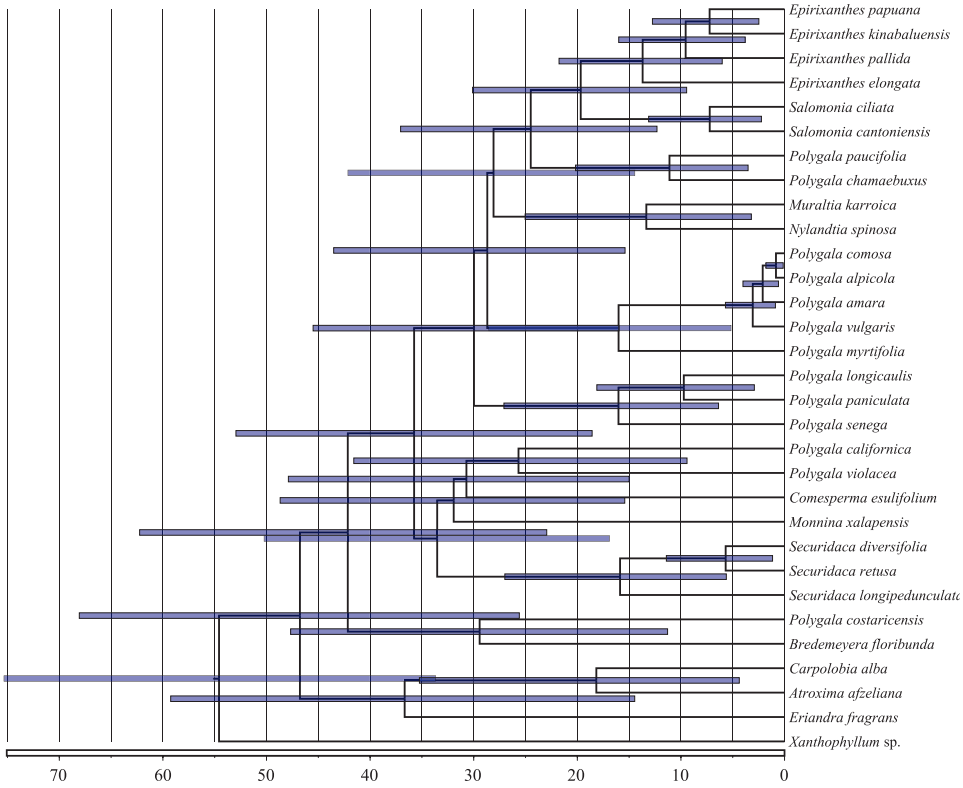


Fig. 14. Maximum clade credibility tree resulting from the divergence time estimation of Polygalaceae inferred using the uncorrelated lognormal relaxed clock method, based on the combined ITS and *matK* data sets. A single secondary calibration point was used to contrain the root of the tree, assuming a Yule process of speciation (see text). The scale bar and values in the figure, represents time (million years ago; Ma), bars around nodes represent 95% confidence intervals.

current taxon sampling. Both genera form a clade (100% BS and 100% PP) that is sister to two species of *Polygala* (*P. chamaebuxus* and *P. pauciflora*). Our results furthermore indicate that within *Epirixanthes*, *E. pallida*, and *E. elongata* form monophyletic groups (both 100% BS and 100% PP). *Epirixanthes papuana* and *E. kinabaluensis* are each represented by a single specimen and form a clade (77% BS and 98% PP). *Epirixanthes pallida*, *E. papuana*, and *E. kinabaluensis* form a clade (100% BS and 100% PP) that is sister to *E. elongata*. The separate analyses of each marker indicate slight topological differences within *Epirixanthes* and more pronounced differences in the rest of the tree, but these conflicting clades never receive strong support (see online Appendix B Figs. S14 and S15). The position of *E. papuana* and *E. kinabaluensis* among the other species is better resolved in the results from the ITS data set (i.e., the early diversification events in *Epirixanthes*), whereas the monophyly of the clade consisting of the two species is only recovered in the *matK* data set. These complementary signals have resulted in well-supported

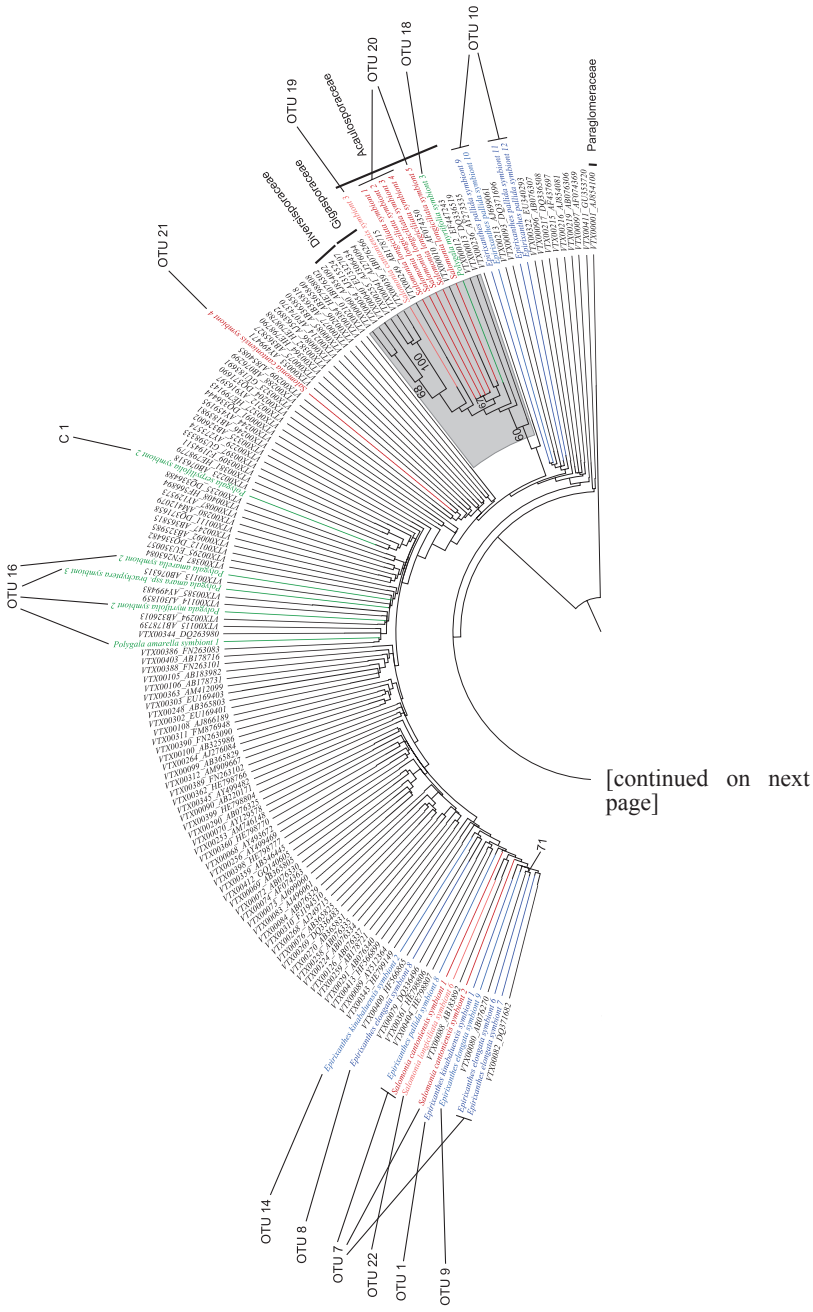
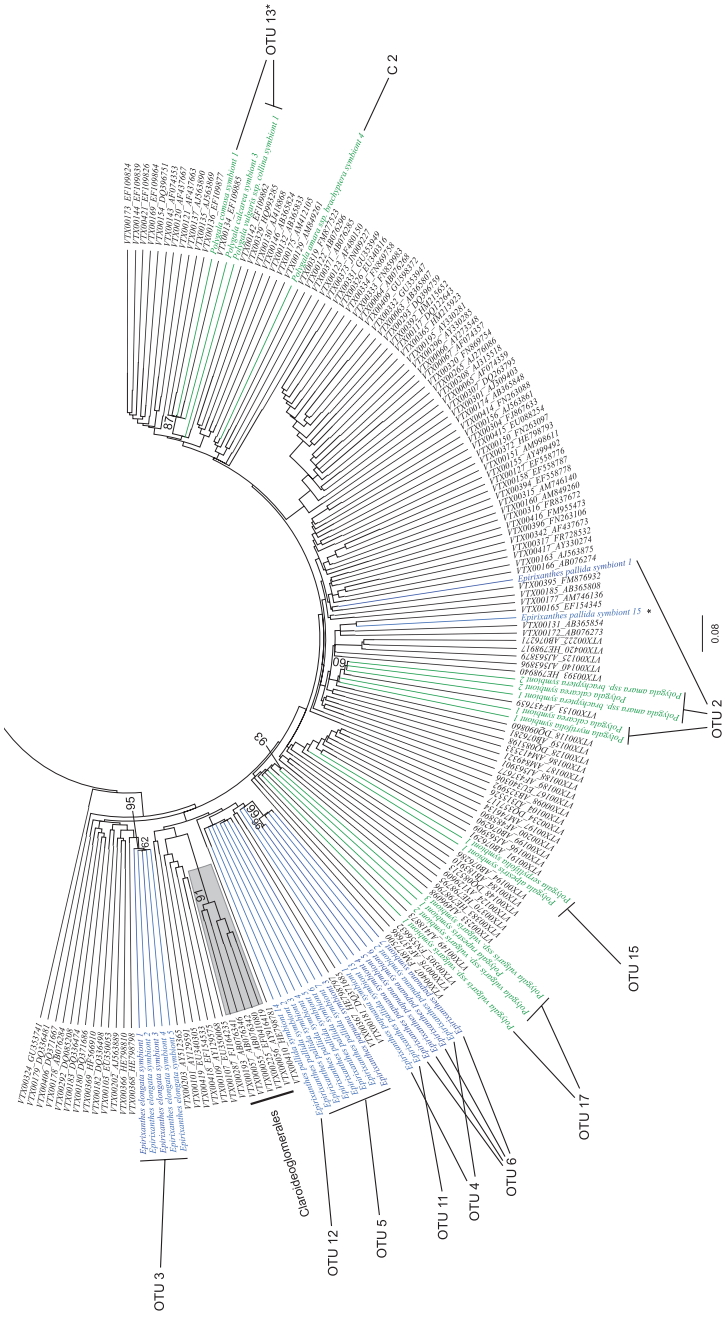


Fig. 15. Phylogeny of arbuscular mycorrhizal fungi based on a maximum likelihood (ML) analysis of a data set consisting of a 500-bp fragment 18S rDNA. Most data are from Glomeraceae, the gray-shaded areas indicate different families. New fungal sequence data for symbionts of *Epirixanthes* (blue) and *Salomonia* (red) are included. Dark colors indicate fungal sequences obtained using cloning, light colors indicate fungal sequences obtained from direct Sanger sequencing. The remaining sequences are obtained from Rath et al. (2013) (*Polygala*) and the Maarj*AM* database (Öpik et al., 2010). Virtual

[Fig. 15 continued]



taxon numbers and GenBank accession numbers are given for the latter. GenBank accession numbers for the remaining fungal sequences can be found in Appendix A.3. Only bootstrap support values of 60% or higher were given for clades containing symbionts of *Epirixanthes*, *Salomonia* or *Polygala.* Corresponding operational taxonomic units (OTUs) are indicated for each fungal sequence, C 1 and C 2 form chimeric sequences which could not be assigned to a single OTU unambiguously. The scale bar indicates the number of substitutions per site.

relationships among *Epirixanthes* species based on the combined data set (Fig. 13). Only the clade formed by *E. papuana* and *E. kinabaluensis* receives moderate bootstrap support, probably resulting from the ITS data.

The clade containing *Epirixanthes*, *Salomonia*, and *Polygala* (96% BS and 100% PP) forms a polytomy with two clades of *Polygala* (96% BS and 100% PP, and 100% BS and 100% PP, respectively) and a clade consisting of *Nylandtia* and *Muraltia* (100% BS and 100% PP). Thus, *Polygala* is paraphyletic. Moreover, the genus *Securidaca* is monophyletic (100% BS and 100% PP), and the tribe furthermore consists of the genera *Comesperma* and *Monnina* (both represented by a single

Table 5. Operational taxonomic units (OTUs), target sequences and VTX affinities.

OTU	Target sequence	Closest hit	Max. % identity (>97%)
1	*Epirixanthes kinabaluensis* symbiont 1	VTX00080	98
2	*Epirixanthes pallida* symbiont 1	VTX00163	98
3	*Epirixanthes elongata* symbiont 5	VTX00368	97
4	*Epirixanthes papuana* symbiont 1	None	
5	*Epirixanthes pallida* symbiont 6	None	
6	*Epirixanthes papuana* symbiont 6	None	
7	*Epirixanthes pallida* symbiont 8	VTX00080	99
8	*Epirixanthes elongata* symbiont 8	None	
9	*Epirixanthes elongata* symbiont 9	VTX00080	97
10	*Epirixanthes pallida* symbiont 12	VTX00093/ VTX00096	97
11	*Epirixanthes pallida* symbiont 13	None	
12	*Epirixanthes pallida* symbiont 14	None	
13	*Polygala comosa* symbiont 1	VTX00135	99
14	*Epirixanthes kinabaluensis* symbiont 2	None	
15	*Polygala alpestris* symbiont 1	VTX00194	100
16	*Polygala amarella* symbiont 1	VTX00115	100
17	*Polygala rupestris* symbiont 1	VTX00149	99
18	*Polygala myrtifolia* symbiont 3	VTX00228	99
19	*Salomonia cantoniensis* symbiont 3	VTX00227	99
20	*Salomonia longiciliata* symbiont 3	VTX00231	99
21	*Salomonia cantoniensis* symbiont 4	VTX00219	98
22	*Salomonia longiciliata* symbiont 6	None	

Table 6. Faith's phylogenetic diversity (PD) scores and their standardized effect sizes (ses.pd) with corresponding *P* values.

Host taxon	No. of AMF sequences	PD	ses.pd	P
Epirixanthes elongata	9	0.48366942	-2.38390807	0.013
Epirixanthes kinabaluensis	2	0.09214911	-1.229551031	0.105
Epirixanthes pallida	15	0.63235641	-3.252918263	0.002
Epirixanthes papuana	7	0.13364764	-4.083472002	0.001
Epirixanthes Total	33	0.98121605	-5.132670632	0.001
Salomonia cantoniensis	4	0.51746465	0.288151921	0.623
Salomonia cantoniensis (Glomeraceae only)	3	0.23563546	-1.172437438	0.178
Salomonia longiciliata	6	0.51530733	-0.888843296	0.184
Salomonia longiciliata (Acaulosporaceae only)	5	0.06487209	-3.864454058	0.001
Salomonia Total	10	0.66123171	-1.594799903	0.084
Polygala alpestris	1	na	na	na
Polygala amara	4	0.36493948	-0.804598192	0.207
Polygala amarella	2	0.02183137	-1.751870053	0.033
Polygala calcarea	3	0.12432341	-1.875023012	0.032
Polygala comosa	1	na	na	na
Polygala myrtifolia	3	0.5373421	1.24807715	0.915
Polygala myrtifolia (Glomeraceae only)	2	0.30147603	0.410710796	0.6015
Polygala rupestris	1	na	na	na
Polygala serpyllifolia	2	0.3776225	0.947202035	0.7365
Polygala vulgaris	4	0.10125083	-2.71228029	0.001
Polygala Total	21	0.92460998	-2.545471824	0.02

specimen). The tribes Polygaleae (27 spp. included in this study, 70% BS and 95% PP) and Carpolobieae (2 spp. included in this study, 97% BS and 100% PP) are recovered as monophyletic groups in the current study, whereas Xanthophylleae and Moutabeae are each represented by a single specimen. Carpolobieae and Moutabieae are sister to each other at the current limited taxon sampling but weakly supported (55% BS and 80% PP), and both are sister to Polygaleae at very weak support (53% BS and 75% PP). The clade consisting of Carbolobieae, Moutabieae, and Polygaleae is well supported and sister to Xanthophylleae (100% BS and 100% PP), providing strong support for the crown group of Polygalaceae, excluding the outgroups from Fabaceae.

4.4.2. Estimation of divergence times

The divergence time estimation of the Polygalaceae data set has resulted in a stem age for *Epirixanthes* of 20 Ma with a 95% confidence interval between 9 and 30 Ma and a crown age of 14 Ma with a 95% confidence interval between 6 and 22 Ma (Fig. 14).

4.4.3. Phylogenetic reconstruction and diversity measurements of AMF

Figure 15 shows the highest likelihood tree resulting from the analysis of partial 18S rDNA extracted from the roots of *Epirixanthes* and amplified using both direct Sanger sequencing and cloning (–lnL = 15 325.3815). For a total of eight plant collections representing all *Epirixanthes* species, several fungal 18S rDNA sequences were obtained by direct sequencing from root tip extractions: 11 for *E. pallida* (obtained from four specimens); two for *E. papuana* (obtained from a single specimen); one for *E. kinabaluensis*; and four for *E. elongata* (obtained from two specimens). For a subset of these accessions, additional 18S rDNA sequences were obtained by cloning; four for *Epirixanthes pallida*, five for *E. papuana*, one for *E. kinabaluensis*, and five for *E. elongata*. For the latter species, two different DNA extractions from the same specimen were used for cloning, resulting in two and three sequences, respectively. For each of the other species, all cloning products result from a single DNA extraction. Two of the obtained fungal sequences of *E. elongata* resulted from the cloning products amplified using the primer combination NS1-EF3 (*E. elongata* symbionts 1 and 3, see Appendix A.3); the rest of the sequences were obtained from the four cloning products using primer combination NS31/AM1. The remaining three of the eight cloning products yielded either poor quality or Ascomycota and/or Basidiomycota data. These sequences all resulted from primer combination NS1 and EF3 (results not shown). The described 18S rDNA data set excluded five full-length duplicates, one of which was extracted from a different species. Six clones of *Salomonia longiciliata* (i1) were obtained, two of which showed 100% sequence similarity. Only a single additional clone was obtained from *Salomonia longiciliata* (i2). DNA extraction of *Salomonia cantoniensis* (i2) yielded five clones, two of which showed 100% sequence similarity.

The arbuscular mycorrhizal symbionts of *Epirixanthes* cluster in 14 OTUs (Fig. 15, OTUs 1–14), of which a single OTU also contains AMF of *Salomonia* and two OTUs also contain AMF of *Polygala*. (See Appendix A.4 for full details on OTU assignment). One of the latter OTUs (OTU 13) contains a single AMF of *Epirixanthes* with a sequence length of only 250 bp, and another OTU of *Epirixanthes* (OTU 14) consists of a single fungal sequence with a length of 372 bp. All other mycorrhizal symbionts obtained from Polygalaceae have sequence lengths between 417 bp and 506 bp. Fourteen OTUs have >97% sequence similarity with Virtual Taxa from the Maarj*AM* database (Öpik et al., 2010), as resulting from BLAST searches of the target

sequences of each OTU (Table 5). The remaining eight OTUs have lower sequence similarities and form potential new Virtual Taxa. Although most OTUs consist of AMF obtained from a single species, some contain AMF from several species. Moreover, many OTUs contain fungal sequences from *Epirixanthes* resulting from both direct sequencing and cloning. Furthermore, the remaining AMF of *Salomonia* and *Polygala* cluster in four OTUs, all of them exclusively containing AMF either obtained from *Salomonia* or *Polygala*. Two chimeric AMF sequences (C 1 and C 2) were obtained from *Polygala*, which could not be assigned to a single OTU.

The phylogenetic reconstruction of these fungi in a Glomeromycota framework based on the Maarj*AM* database (Öpik et al., 2010) indicates that all AMF obtained from *Epirixanthes* belong to Glomeraceae (the former *Glomus* Group A) (Fig. 15). The majority of the analyzed mycorrhizal symbionts of *Polygala* also belong to Glomeraceae, though a single fungal sequence was found to belong to Acaulosporaceae. For *Salomonia longiciliata*, however, the majority of the AMF are recovered in Acaulosporaceae, with only a single exception from Glomeraceae. However, most AMF of *S. cantoniensis* emerge in Glomeraceae, and a single fungal sequence from this species belongs to Acaulosporaceae. Support values are given for clades containing mycorrhizal symbionts from either of the studied genera and receiving over 60% BS. These are generally very low; only a single clade of AMF of *Epirixanthes* (consisting of AMF of *E. elongata*) is well supported, as well as the Claroideoglomerales, Gigasporaceae, and some clades consisting of AMF of *Polygala*, mostly together with virtual taxa. However, although all non-Glomeraceae families/orders are recovered as clades, their position is nested within Glomeraceae (except the outgroup of Paraglomeraceae). Furthermore, the recovered OTUs roughly coincide with the clades in this phylogeny. Fungi obtained from *Polygala* show a pattern highly similar to the findings of Rath et al. (2013), with a single fungal sequence belonging to Acaulosporaceae (a symbiont of *P. myrtifolia*).

4.4.4. Phylogenetic diversity measurements of AMF

Phylogenetic diversity values and their standardized effect size values (SES_{PD}) are summarized in Table 6. Fungi of *Epirixanthes* have a PD of 0.48 (*E. elongata*), 0.09 (*E. kinabaluensis*), 0.63 (*E. pallida*), and 0.13 (*E. papuana*). However, all except *E. kinabaluensis* have significantly negative SES_{PD} values; -2.38 ($P = 0.013$) for *E. elongata*, -3.25 ($P = 0.002$) for *E. pallida*, and -4.08 ($P = 0.001$) for *E. papuana*. These results suggest the AMF recovered within each of these species are more closely related to each other than expected under the applied null-model based randomization. Fungi of both *Salomonia cantoniensis* and *S. longiciliata* have a PD of 0.52. In the former species, predominantly AMF of Glomeraceae are found, with only a single exception from Acaulosporaceae. The opposite pattern is found in the latter species, which grows predominantly in symbiosis with Acaulosporaceae, with a single exception from Glomeraceae. Exclusion of these exceptional fungal

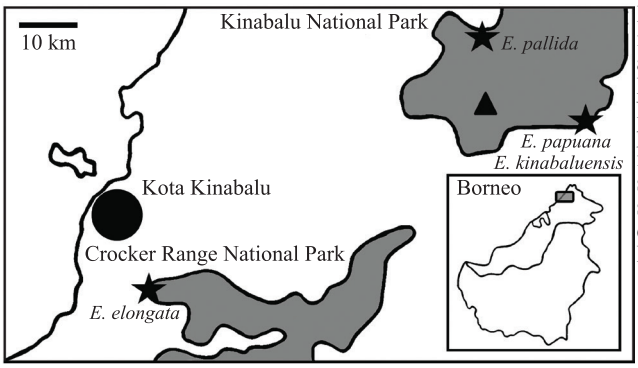


Fig. 16. Map of the collection sites in Sabah, Malaysian Borneo. Gray areas indicate Kinabalu National Park (top) and Crocker Range National Park (bottom), where the plants were collected. Black stars indicate collection sites for each species; black triangle indicates the summit of Mount Kinabalu; black dot indicates the urban area of Kota Kinabalu.

sequences results in a strong drop in PD values in both species, 0.24 in *S. cantoniensis* and 0.06 in *S. longiciliata*. The latter value has a significantly negative SES_{PD} of -3.86 ($P = 0.001$), also indicating these fungi are more closely related to each other than expected under the performed randomization. The AMF of the nine species of *Polygala* have PD values varying between 0.02 for AMF of *P. amarella* and 0.54 for AMF of *P. myrtifolia*. This high PD value for AMF of *P. myrtifolia* results from a single Acaulosporaceae sequence (the only one found in *Polygala*), and removal of it results in a drop of the PD value to 0.30. Only in *P. vulgaris* a significantly negative SES_{PD} value is recovered (-2.71, $P = 0.001$), also indicating a closer relationship than expected under randomization. For three species (*P. alpestris*, *P. comosa*, and *P. rupestris*) only a single fungal sequence was included, which implies no PD values could be calculated. Although no significantly positive SES_{PD} values are recovered, the number of fungal sequences per species was mostly lower in *Polygala* as compared with *Epirixanthes* and *Salomonia*. Therefore, the PD values in *Polygala* may be (slightly) underestimated, suggesting that most phylogenetic diversity (i.e., all values above 0.50 without significant SES_{PD} values) is found in *Salomonia* due to the recovery of both Glomeraceae and Acaulosporaceae from both its species. The same is found in *P. myrtifolia*, but not in other species of *Polygala*.

4.5. Discussion

4.5.1. Phylogeny of *Epirixanthes*

Epirixanthes forms a strongly supported monophyletic genus at current taxon sampling. The placement of the genus in tribe Polygaleae is in accordance with earlier observations and suggestions (e.g., Van der Meijden, 1988; Bello et al., 2012). The sister-group relation with *Salomonia* (of which both accepted species by Van der Meijden (1988) are included in this study) supports the proposed close affinity of the genera (Chodat, 1896; Hutchinson, 1967; Van der Meijden, 1988). Although they form a monophyletic group, recognition of two separate genera is warranted, because each genus is monophyletic. Moreover, clear differences in habit (as a result

of the different modes of life of the genera; mycoheterotrophic vs. autotrophic) and floral morphology are reported (Van der Meijden, 1988).

The intrageneric relationships of *Epirixanthes* are in accordance with the species delimitations, suggesting that the used DNA markers are effective DNA barcodes for species identification in this genus. The phylogenetic relationships between the species show remarkable geographic structure: all specimens of *E. pallida*, *E. kinabaluensis*, and *E. papuana* were found in Kinabalu National Park and form a clade (see Fig. 16). The former species is sister to the rest and was exclusively found in one locality at the northern slope of Mount Kinabalu, whereas *E. kinabaluensis* and *E. papuana*, each represented by a single specimen, form a moderately supported clade corresponding to a locality on the eastern slope. *Epirixanthes elongata* forms a monophyletic group as sister to the rest of the genus and was the only species found in Crocker Range National Park. A single specimen from Sumatra, represented solely by an ITS sequence, is confidently placed in the clade forming *E. elongata*. This evidence suggests that the found geographical structure is a sampling artifact, possibly due to different microclimates at each site resulting in different species occurrence (or abundance) at the time of collecting. It is unlikely that the found species diversity resulted from local speciation in Kinabalu National Park or Crocker Range National Park, as the studied species all have wider distributions in Southeast Asia. Inclusion of more specimens of the same species from elsewhere in Southeast Asia will probably result in monophyletic species rather than geographically structured clades. The observation of four *Epirixanthes* species from only two adjacent National Parks in Sabah, Borneo, supports that Borneo is a hotspot of *Epirixanthes* diversity. However, including more specimens from different areas throughout the distribution range of *Epirixanthes* would be informative.

The close relationship between *Epirixanthes* and *Salomonia* is supported by a synapomorphy (i.e., three antesepalous stamen primordia) as well as some additional morphological similarities, such as a spike-like terminal inflorescence (Van der Meijden, 1988). Both genera are confidently placed in Polygalaceae, but our data provides limited resolution in the rest of the family (Fig. 13; Appendix B Figs. S14 and S15). Other studies (e.g., Persson, 2001; Forest et al., 2007; Bello et al., 2009, 2012) address phylogenetic relationships at the family and order levels in more detail.

4.5.2. Estimation of divergence times

The divergence time estimation of *Epirixanthes* suggests a Miocene-Oligocene stem age and a Miocene crown age of the genus. Diversification of *Epirixanthes* coincides with much tectonic activity and climatic changes in Southeast Asia (Hall, 2009; De Bruyn et al., 2014). These complex environmental dynamics may have triggered speciation in the genus. In particular, the mid-Miocene Climatic Optimum, when Asian rainforests reached their maximum northward extent (Böhme, 2003;

De Bruyn et al., 2014), is likely to have had a major influence on *Epirixanthes* evolution. The current wide and partly overlapping distribution of many of the species of *Epirixanthes* might be the result of the many possible land connections between Borneo, the Southeast Asian mainland and other parts of the region during the last 30 Ma (Hall, 2009; De Bruyn et al., 2014). Dispersal over water is a less likely mechanism, as seed morphology of *Epirixanthes* suggests dispersal by ants (Van der Meijden, 1988). Inclusion of conspecific *Epirixanthes* specimens from elsewhere in Asia would be highly informative to describe the biogeographic history of *Epirixanthes* in more detail. Moreover, our divergence time estimation is only based on four of the six described species of *Epirixanthes*. The species that were not sampled, however, are not likely to have a profound effect on the divergence age estimates, as both species are thought to be very closely related to the species represented here (Van der Meijden, 1988; Pendry, 2010). Another limitation of the presented divergence time estimation is the limited calibration information available. Although we assume the secondary calibration point from Forest et al. (2007) provides a reliable crown age of Polygalaceae, the lack of primary (fossil) calibration points used in this study implies that the result should be interpreted as a rough approximation of the actual divergence times.

4.5.3. Phylogeny of AMF

AMF extracted from the roots of *Epirixanthes* species are all from Glomeraceae (the former *Glomus* Group A). On the basis of morphology, Imhof (2007) also found AMF in the roots of *Epirixanthes elongata* and *E. papuana*, which he interpreted as *Paris*-type AMF (Gallaud, 1905). Although this type is formed by Glomeromycota (Smith and Read, 2008) its presence was not assessed in this study. In all species of *Epirixanthes*, more than one OTU of AMF was amplified, and distinct species were often found to grow on the same AMF. This pattern of symbiotic AMF diversity suggests that species of *Epirixanthes* do not grow on species-specific clades of AMF. Fungi of Ascomycota and Basidiomycota were amplified from the roots of *Epirixanthes* as well, but these resulted from amplifications with primer combination NS1 and EF3, none of which is specific for Glomeromycota (White et al., 1990; Smit et al., 1999). The BLAST search revealed those fungi are plant pathogens and/or other symbionts that were assumed to be no part of the mycorrhizal interaction and therefore not further analyzed. We assume that the mycorrhizal symbionts of *Epirixanthes* only consist of the amplified AMF.

This study confidently indicates a close relationship between *Epirixanthes*, *Salomonia*, and *Polygala*. Due to their different type of symbiosis with their mycorrhizal fungi (i.e., supposedly parasitic in *Epirixanthes* and mutualistic in *Salomonia* and *Polygala*), the comparison between AMF from the roots of the genera provides information on the evolution of mycoheterotrophy in Polygalaceae. About half of the plants from which AMF were isolated by Rath et al. (2013) were

also included in the phylogenetic reconstruction presented above. The study of Rath et al. (2013) indicates these closely related *Polygala* species are predominantly symbiotically connected to AMF of Glomeraceae. Comparison between the symbiotic AMF from the roots of Bornean *Epirixanthes* and the AMF from the roots of Thai *Salomonia* and the European *Polygala* species from Rath et al. (2013) indicates the genera are symbiotically linked to a similar and slightly overlapping range of AMF of Glomeraceae (Fig. 15). Three distantly related Glomeraceae OTUs (OTU 7, OTU 21, and OTU 22) were recovered from the roots of *Salomonia* species, one of which (OTU 7) also contains AMF of *Epirixanthes pallida* and *E. elongata*. This suggests some overlap in symbiotic fungi, though in general *Epirixanthes* and *Salomonia* are linked to a different set of AMF. Moreover, two OTUs (OTU 2 and OTU 13) containing several AMF sequences from *Polygala* species also contain a single fungal sequence from *Epirixanthes pallida*. However, one of those sequences from *Epirixanthes* (*E. pallida* symbiont 15) is rather short (see Appendix A.4) and its OTU clustering is therefore less reliable. However, a single fungal sequence of Acaulosporaceae was obtained from both *Polygala myrtifolia* and *Salomonia cantoniensis*. *Salomonia longiciliata* is predominantly associated with this family of AMF, despite its close relationship with *Epirixanthes*. Apart from some overlapping AMF between *Epirixanthes* and the other genera, 11 OTUs are unique for *Epirixanthes*, four are unique for *Salomonia* and four are unique for *Polygala*. These results suggest that *Epirixanthes* is symbiotically linked to a wide range of AMF, exclusively from Glomeraceae, many of which are not associated to either *Salomonia* or *Polygala*. However, due to limited taxon sampling and low support values in the AMF phylogeny, the patterns emerging from these data should be interpreted with care. Moreover, as the AMF amplification workflow was different between the two studied plant genera, as well as the polymerase and primer combinations used, the resulting patterns might be biased. This may also affect the comparison to previous studies (e.g., Rath et al., 2013; Merckx et al., 2012). However, primers designed specifically for AMF were used in both genera (NS31 [Simon, 1996] and AM1 [Helgason et al., 1998]), supplemented by less-specific primers (NS1 [White et al., 1990] and EF3 [Smit et al., 1999] for Epirixanthes; GeoA1 and ART4 [Schwarzott and Schüßler, 2001] for *Salomonia*) to compensate a potential bias by NS31 and AM1. We therefore assume that a careful comparison can be made despite the different approaches and that the obtained sequences of mycorrhizal symbionts of both genera provide a representative sample of the actual mycorrhizal community in the roots of these plants. Moreover, comparisons with different studies should also be made carefully, under the assumption that these studies provide a similar representative sample of AMF.

4.5.4. Phylogenetic diversity measurements of AMF

The wider range of AMF symbionts of *Salomonia* and *Polygala* as compared with *Epirixanthes* indicated by the phylogeny of the AMF is corroborated by higher PD

values for the fungi from these genera, particularly from *Salomonia*. SES_{PD} values were calculated to account for the sampling differences between the species (see Table 6). The significantly negative SES_{PD} values for most of the AMF of the *Epirixanthes* species and to a lesser extent for those of *Salomonia* (the Acaulosporaceae fungi of *S. longiciliata*) and *Polygala* (*P. vulgaris*) suggest oversampling, probably leading to an overestimated PD value. These SES_{PD} values suggest the actual PD of these taxa is lower, which particularly supports the finding of a low PD of the AMF from the *Epirixanthes* species as compared with the PD of the AMF of *Salomonia*. Comparison between AMF of *Epirixanthes* and *Polygala* is difficult due to the different number of species and fungal sequences per species. Moreover, sampling of AMF of *Polygala* was done over a relatively large surface area (i.e., Germany and two islands in the Mediterranean Sea) in a temperate region, as compared with the two Bornean rainforest areas where *Epirixanthes* was collected. However, the finding of Acaulosporaceae in a single species of Polygalaceae also indicates higher diversity in this genus as compared with *Epirixanthes*. The results indicate the high PD values in *Salomonia* and *Polygala* are largely the result of their symbiosis with both Glomeraceae and Acaulosporaceae, whereas *Epirixanthes* is only linked to Glomeraceae.

This raises the question of whether species of *Epirixanthes* are specialized on AMF. As *Polygala* species were interpreted as being not relatively selective on AMF (i.e., symbiotically linked to five distinct clades; Rath et al., 2013), the results from this study do not provide a straightforward answer. *Epirixanthes* species could be interpreted as being more selective than those of *Salomonia* and *Polygala*, given the lack of Acaulosporaceae fungi found in its roots and its lower PD value. However, given the wide range of symbiotic fungi extracted from the roots (i.e., 14 OTUs scattered across Glomeraceae) and the limited distribution of the sampled *Epirixanthes* specimens, the results tend to suggest species of *Epirixanthes* are not highly specialized on specific clades of Glomeraceae. As the majority of AMF engaged in mycoheterotrophic interactions are Glomeraceae (Merckx et al., 2012), only plants with a restricted clade of symbiotic Glomeraceae are interpreted as having high levels of specialization (e.g., *Arachnitis uniflora*, Corsiaceae [Bidartondo et al., 2002]; *Afrothismia* spp., Thismiaceae [Merckx and Bidartondo, 2008]). Lack of specificity however, is not a rare phenomenon in mycoheterotrophic plants as it is found in species of *Aphyllorchis* (Orchidaceae), *Cephalanthera* (Orchidaceae), *Pyrola* (Ericaceae), *Sciaphila* (Triuridaceae) and *Campylosiphon* (Burmanniaceae) (Hynson and Bruns, 2009; Roy et al., 2009; Merckx et al., 2012). Concluding, our results suggest that species of *Epirixanthes* are more specialized than their chlorophyllous relatives, but the species are not highly specialized as compared with many other mycoheterotrophic plants. More research, however, is needed to elucidate specificity in Polygalaceae further, especially using uniform sampling and molecular methods.

Epirixanthes appears to have evolved relatively recently toward mycoheterotrophy, as many groups of mycoheterotrophic plants are much older; e.g., Triuridaceae have an estimated mean stem age of 84 Ma (Mennes et al., 2013), and Corsiaceae have an estimated mean stem age of 70 Ma (Mennes et al., 2015a). The relatively wide range of AMF the genus is symbiotically linked to possibly represents a relic of a more generalistic ancestral fungal symbiosis, similar to that of *Polygala* (Rath et al., 2013) or *Salomonia*. However, generalistic fungal symbioses are also observed in other mycohetrotrophic lineages (e.g., *Sciaphila* and *Campylosiphon*; Merckx et al., 2012), suggesting this type of symbiosis might not necessarily represent such an ancestral state. Moreover, the strong morphological similarity to its closest chlorophyllous relative (*Salomonia*) and the relatively short branch lengths, suggest that *Epirixanthes* might not have undergone the same degree of evolutionary change as many other clades of mycoheterotrophic genera, e.g., *Petrosavia*, *Afrothismia*, and *Voyria*. Last, the functional reading frame in the *matK* region for at least some *Epirixanthes* specimens might indicate this plastid region is still functional. As this region is believed to be one of the last to lose its functionality in the transition toward mycoheterotrophy (Barrett and Davis, 2012), this observation further suggests the evolutionary shift toward mycoheterotrophy in *Epirixanthes* is a relatively recent one. These results support the hypothesis that loss of photosynthetic ability in mycoheterotrophic plants is not necessarily correlated with specialization to narrow lineages of fungi (Hynson and Bruns, 2009). The current study shows that *Epirixanthes* forms an excellent system to study the evolution of mycoheterotrophy.

4.6. Acknowledgements

The authors thank M. Schilthuizen, M. Lakim, R. Repin, and S. Sasirat for help and assistance in obtaining plant material and M. Eurlings and Y. Sirichamorn for technical support. They also thank the Alberta Mennega Foundation and the Schure Beijerinck Popping Foundation (now KNAW Fonds Ecologie, Dutch Academy of Sciences) for financial support of the fieldtrip to Malaysian Borneo to collect the Epirixanthes material. V.S.F.T.M. is financially supported by NWO (Veni fellowship 863.11.018).